

Un'ontologia formale per la creazione automatica di proposte formative: il progetto E-VALUE

Luciano Serafini¹, Volha Bryl¹, Sofia Cramerotti², Chiara Ravagni²,

¹Fondazione Bruno Kessler, Via Sommarive 18, 38123 Trento, TN
serafini@fbk.eu , bryl@fbk.eu

²Centro Studi Erickson, Via del Pioppeto 24, 38121 Gardolo, Trento, TN
safia.cramerotti@erickson.it , chiara.ravagni@gmail.com

In questo articolo viene illustrato lo schema di rappresentazione e classificazione dei learning object utilizzato all'interno del progetto E-VALUE. Viene inoltre descritto come questo è stato codificato in un'ontologia in linguaggio OWL e vengono mostrati alcuni esempi di query.

1. Introduzione: il progetto E-VALUE

L'individuazione precoce dei fattori di rischio e la successiva messa in atto di percorsi di intervento riferiti alle principali **difficoltà di apprendimento** è un processo che va strutturato seguendo rigorose procedure e parametri di riferimento, basati sulle più accreditate classificazioni e Linee guida condivise dalla comunità scientifica. In particolare la **Consensus Conference sui DSA** (2009) riporta degli indicatori di rischio che hanno mostrato un buon valore predittivo rispetto alle difficoltà di apprendimento; la loro individuazione dovrebbe permettere successivamente di attivare percorsi adeguati di recupero/potenziamento.

Partendo da queste premesse, il **progetto E-VALUE** intende offrire una gamma di **strumenti di valutazione e di materiali per l'intervento educativo-didattico** nei principali ambiti di apprendimento (in particolare lettura, scrittura e calcolo). La finalità è quella di strutturare un sistema integrato per la valutazione di tali abilità e la gestione del processo di intervento didattico.

1.1 Obiettivi e finalità del progetto

L'obiettivo principale del progetto E-VALUE è progettare, sviluppare e sperimentare un sistema informatico integrato – piattaforma multimediale (costituito da un insieme di applicativi web), che consenta la **valutazione delle abilità di apprendimento e la gestione del processo di intervento didattico** all'interno di diversi contesti educativi. Tale sistema è costruito secondo un modello che integra processi differenti, interconnessi sotto il profilo funzionale: procedure di valutazione standard, valutazione dinamica, raccolta di dati normativi e continua "alimentazione" (aggiornamento della banca dati e

confronto con dati normativi attualmente a disposizione), lettura qualitativa oltre che quantitativa dei dati raccolti, impostazione dei percorsi di intervento, metodologie/materiali didattici direttamente spendibili nella progettazione educativo-didattica quotidiana. Tutti questi processi/attività fanno parte di modelli di valutazione/intervento attuati da anni “direttamente sul campo” in scuole di ogni ordine e grado, e che fanno riferimento alle più accreditate modalità di sperimentazione educativo-didattica: percorsi di ricerca-azione, utilizzo di user-group, procedure di test-retest, uso/confronto tra gruppi sperimentali e gruppi di controllo.

In quest’ottica il **sistema implementato all’interno del Progetto E-VALUE** permette quindi:

- la gestione e l’erogazione di servizi, strumenti di rilevazione e contenuti didattici finalizzati alla valutazione psicoeducativa degli alunni in contesti scolastici;
- la conservazione dei contenuti (materiali per la valutazione, materiali per il recupero e il potenziamento), in formato digitale all’interno di appositi archivi al fine di facilitare la ricerca e l’utilizzo delle risorse;
- la gestione delle fasi (workflow) del processo di valutazione attraverso un sistema (workflow management system) in grado di gestire in modo automatico il passaggio dei dati tra le varie fasi diminuendo i tempi di risposta e aumentando l’efficienza del servizio;
- la standardizzazione dei metodi e degli strumenti utilizzati nelle fasi di valutazione /scoring;
- la rappresentatività dei dati rilevati grazie al progressivo ampliamento della banca dati normativa di riferimento.

1.2 Articolazione del Modello E-VALUE

Il modello E-VALUE si articola in:

Valutazione: scelta dello strumento e modalità di somministrazione; somministrazione delle prove secondo la modalità scelta (individuale/collettiva); raccolta delle risposte e invio dei risultati.

Scoring/Restituzione: elaborazione dati quanti-qualitativa (comunità di esperti); restituzione dei risultati in presenza e a distanza; intervento didattico, e strutturazione del percorso di recupero/potenziamento delle abilità/competenze del singolo o del gruppo classe (esperti, referente scolastico).

Intervento: progettazione del percorso di intervento (metodologie-contenuti) da parte del referente scolastico in collaborazione con l’esperto; intervento in laboratori didattici in cui vengono coinvolti tutti gli alunni della classe; tutoraggio e monitoraggio in itinere delle attività di recupero/potenziamento.

Tutto il percorso è preceduto e accompagnato da una formazione continua: iniziale, in itinere (su esito dei risultati e progettazione delle attività) e da momenti di consulenza/tutoraggio, nonché da un’attenta verifica dell’efficacia del progetto educativo-didattico che si sta attuando.

1.3 Analisi dei materiali

Lo scopo dell'analisi dei materiali di intervento didattico è quello di selezionare le parti, da libri o da CD, che intervengono su un determinato processo cognitivo. Per farlo è stato necessario:

- identificare i processi cognitivi
- analizzare il materiale pubblicato dalle Edizioni Centro Studi Erickson e scomporlo in parti autonome
- trasformare il materiale in modo da trasferirlo dal supporto ottico o cartaceo a una struttura in grado di ricomporlo secondo necessità, online e offline.

Nella struttura tradizionale del libro o del CD, infatti, i materiali sono ordinati secondo un percorso didattico progressivo che coinvolge svariati processi cognitivi. Nella struttura del progetto E-VALUE invece i materiali vengono ricombinati in modo nuovo, divisi in parti con contenuto e processo cognitivo uniforme e utilizzabili in determinate condizioni (ad esempio una fascia d'età, oppure singoli alunni piuttosto che l'intera classe).

2. La catalogazione

Si poneva quindi la necessità di predisporre una forma di catalogazione adatta a descrivere il materiale da ogni punto di vista utile al futuro sistema, per la stesura della quale si è tenuto conto dei seguenti aspetti:

- la necessità di descrivere il materiale analizzato dagli esperti in modo completo, dettagliato e bibliograficamente uniforme;
- la collocazione in una struttura di dati da cui fosse sempre deducibile il materiale d'origine e i materiali che da qualunque punto di vista (editoriale, contenutistico, informatico) erano ad esso correlati;
- il bisogno di organizzare la descrizione in una lista di "tag" tali da essere compresi e interpretati automaticamente da un sistema di decisione automatica o da un motore di ricerca per il reperimento di materiale da parte di formatori e collaboratori.

2.1. Descrizione degli oggetti

Si è scelto quindi di produrre una struttura di catalogazione dettagliata che combinasse i campi catalografici del Dublin Core con campi appositamente creati per rispettare le necessità del gruppo di lavoro e del progetto.

I 15 campi del Dublin Core si sono rivelati sufficientemente generici da poter appoggiare ad essi un ampio numero di qualificatori per un maggiore dettaglio nella descrizione. Inoltre sono stati aggiunte alcune categorie di tag più

propriamente amministrative/gestionali, per poter trattare il materiale anche dal punto di vista commerciale e dei contratti con gli autori.

I tag si possono quindi raggruppare in:

- estratti dal Dublin Core
- descrittivi
- gestionali.

La descrizione rispecchia una struttura a faccette spuria, considerando che i campi di tipo contenutistico (quelli che riguardano i processi cognitivi, le funzioni cognitive associate e il livello di competenza) permettono l'accesso ai materiali tramite *query* diverse.

2.2 I vocabolari controllati

Viste le difficoltà nella gestione delle forme di catalogazione precedenti si è deciso di corredare la tabella con una serie di dizionari controllati. Ad ogni campo (tranne i campi autore/titolo e alcuni campi testuali liberi) è stata affiancata una lista di parole che compongono una scelta obbligata per il catalogatore. Questo è necessario per evitare la scelta di termini simili o dello stesso concetto ma in un'altra forma grammaticale (singolare o plurale, aggettivo o sostantivo), fatto che renderebbe molto difficoltoso il *mapping* dei materiali per il DSS. Questi dizionari sono stati concordati durante le riunioni del gruppo di lavoro.

Nel caso dei campi riguardanti il contenuto (area tematica, abilità generale, processo, sottoprocesso) per poter descrivere in modo dettagliato il materiale è stato compilato un amplissimo dizionario controllato con una struttura ad albero su quattro livelli che copre gli ambiti di: matematica, letto-scrittura, metafonologia e abilità linguistiche. Per la stesura del dizionario sono state necessarie periodiche riunioni con esperti delle varie tematiche.

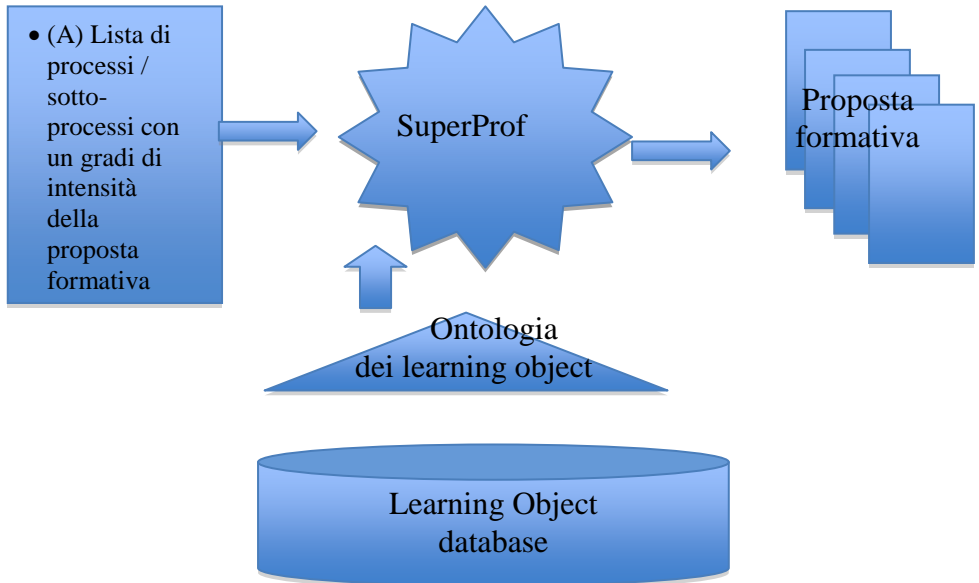
Per riuscire a descrivere l'ambito in modo adeguato si è utilizzato il metodo del brainstorming con esperti, seguito da fasi di test in cui le varie combinazioni di metadati sono state provate nella loro efficacia e modificate quando si rivelavano poco funzionali o incomplete.

3. L'ontologia

Prima di introdurre specificatamente l'ontologia associata al progetto E-VALUE, è importante capire quale sia l'utilizzo di questa ontologia all'interno del progetto. Per questo, nella prima parte di questo paragrafo descriveremo a un livello molto generale, il sistema di supporto alle decisioni per la generazione di proposte formative che verrà realizzato all'interno del progetto.

3.1 Il sistema di supporto alle decisioni

Il sistema di supporto alle decisioni che verrà realizzato in E-VALUE, denominato SuperProf (**S**istema di **S**upporto alle decisioni **PER** la creazione di **PRO**poste **FOR**mativa) ha la funzionalità principale di generare, a partire dalla specifica di una necessità formativa, data in termini di processi/sottoprocessi formativi, una proposta formativa, composta da una insieme di *learning object*. Una visione ad alto livello del sistema di supporto alle decisioni SuperProf è schematizzata nella seguente figura



Il sistema di supporto alle decisioni, SuperProf prende in input una lista di processi/sotto-processi, ognuno dei quali è associato a una serie di parametri che indicano l'intensità dell'intervento formativo necessario in quel processo. In aggiunta a queste informazioni, al sistema vengono forniti una serie di parametri, quali ad esempio: numero totale di *learning object* di cui è composta la proposta formativa, percentuali di mix tra diverse tipologie di *learning object*, che verranno utilizzate da SuperProf per modulare la proposta formativa.

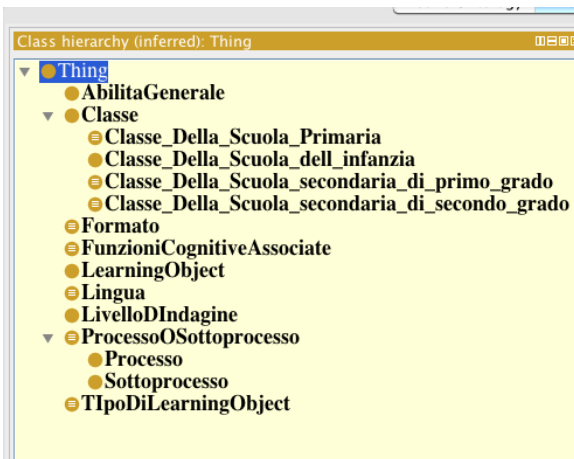
A fronte di questo tipo di input il DSS SuperProf, in una prima fase, seleziona tra tutti i *learning object* disponibili nel catalogo Erickson quelli rilevanti per i processi/sottoprocessi specificati nell'input (A), e successivamente sceglie tra quelli estratti dal database un sottoinsieme che soddisfi nel miglior modo possibile i criteri aggiuntivi (B) specificati nell'input. La classificazione dei *learning object* ha un ruolo fondamentale soprattutto nella prima fase di selezione. Infatti in questa fase, a fronte di una richiesta di attività formativa su un insieme di processi/sottoprocessi, verranno selezionati i *learning object* che sono stati classificati con i relativi tags, o con tags equivalenti. È in questa fase che la codifica della classificazione dei LO in un'ontologia diventa

fondamentale, in quanto permette di realizzare la selezione/estrazione degli LO rilevanti in modo semplice, flessibile ed efficiente.

3.2 L'ontologia di E-VALUE

Il termine Ontologia può avere molti significati ma nell'ambito dei sistemi per la gestione della conoscenza questo termine viene ormai costantemente associato alla definizione data da Gruber (1993) secondo la quale un' ontologia è la caratterizzazione di una concettualizzazione

condivisa. Il linguaggio standard per la caratterizzazione delle ontologie che si è affermato negli ultimi anni è OWL (Ontology Web Language). Questo linguaggio, la cui semantica formale è basata sulla logica descrittiva, permette di descrivere un' ontologia tramite due componenti principali, la T-box (*terminological box*) che contiene la gerarchia delle



classi-sottoclassi dell'ontologia e le relazioni che possono esistere tra le varie classi, e la A-Box (*assertional box*) che contiene gli oggetti dell'ontologia, la loro classificazione (cioè l'appartenenza di un oggetto ad una classe/sotto-classe) e le relazioni tra gli oggetti. L'ontologia di E-VALUE può essere visualizzata utilizzando l'editor di Ontologie *Protegé* disponibile al sito <http://protege.stanford.edu/>

La gerarchia delle classi dell'ontologia di E-VALUE, mostrata nella seguente figura, è composta dalle seguenti classi principali

AbilitaGenerale è la classe che contiene tutte le abilità generali, associate ai vari livelli di indagine, individuate nel modello di classificazione di E-VALUE. Esempi di elementi di questa classe sono: "*Letto-scrittura, decodifica*", "*Letto-scrittura, comprensione*" e "*Letto-scrittura, compilazione*".

Classe questa classe contiene tutte le classi dalla scuola dell'infanzia alla scuola secondaria di secondo grado, organizzate in sottoclassi, per ogni periodo scolastico omogeneo. Esempi di istanze di questa classe sono: *1P, 2P ... 5P* (le cinque classi della scuola primaria) *1S1, 1S2, 1S3* (le tre classi della scuola secondaria di primo grado).

Formato è la classe che contiene tutti i possibili tipi di formato dei *learning object*. Gli elementi appartenenti a questa classe sono: *cartaceo, ebook, materiale*.

FunzioniCognitiveAssociate è la classe che contiene tutte le funzioni cognitive secondo la classificazione introdotta nella sezione precedente. Alcuni

esempi di elementi di questa classe sono: *FCA_Attenzione_verbale*, *FCA_Abilita_visuo_spaziali*, *FCA_Attenzione_uditiva*, *FCA_Attenzione_visiva*. **LearningObject** è la classe che contiene tutti i *learning object* presenti nel catalogo. Ogni *learning object* ha un identificativo unico.

Lingua è la classe che contiene l'insieme delle lingue nelle quali sono scritti i *learning object*.

LivelloDindagine è la classe che contiene i diversi livelli di indagine individuati nello schema di E-VALUE. Alcuni esempi di livelli di indagine sono: *Metafonologia*, *Letto-scrittura*, *Abilità linguistiche*, *Matematica-Geometria*,

ProcessoOSottoprocesso è la classe che contiene tutti i processi raggruppati nei due livelli, processi e sottoprocessi. Esempi di processi sono: *Metafonologia-globale-rime* e *Metafonologia-globale-sillabe*; i relativi sottoprocessi sono *riconoscimento* e *produzione* (sottoprocessi del primo) e *fusione orale* e *fusione orale con supporto visivo* (sottoprocessi del secondo)

TipoDiLearningObject è la classe che contiene tutte le tipologie di *learning object* quali ad esempio: *Libro*, *TLearningObject*, *ILearningObject*, *Articolo_di_rivista*, *Materiale_stand-alone* e *Flash*

Il secondo gruppo di elementi della T-box dell'ontologia di E-VALUE sono le relazioni. In generale le relazioni di un'ontologia sono relazioni binarie che collegano gli oggetti tra loro. Una vista generale delle relazioni (chiamate anche *ObjectProperties*) presenti nell'ontologia di E-VALUE, e fornita nella figura a fianco. Nel seguito ne diamo una descrizione più dettagliata, raggruppandole in due categorie. Nella prima troviamo le relazioni che servono a caratterizzare ciascun *learning object*. Sono relazioni che associano ad ognuno una serie di attributi che lo caratterizzano, come ad esempio: il tipo, il range di classi a cui è indirizzato, la lingua, il formato o il tipo.

haClasseDa e **haClasseA** sono due relazioni che associano ad ogni Learning Object il range di classi a cui esso è indirizzato. Se per esempio un *learning object* LO è indicato per la classe 1°, 2° e 3° elementare, la relazione *haClasseDa* associerà 1P (1° elementare) al LO, e la relazione *haClasseA* associerà 3P a LO.

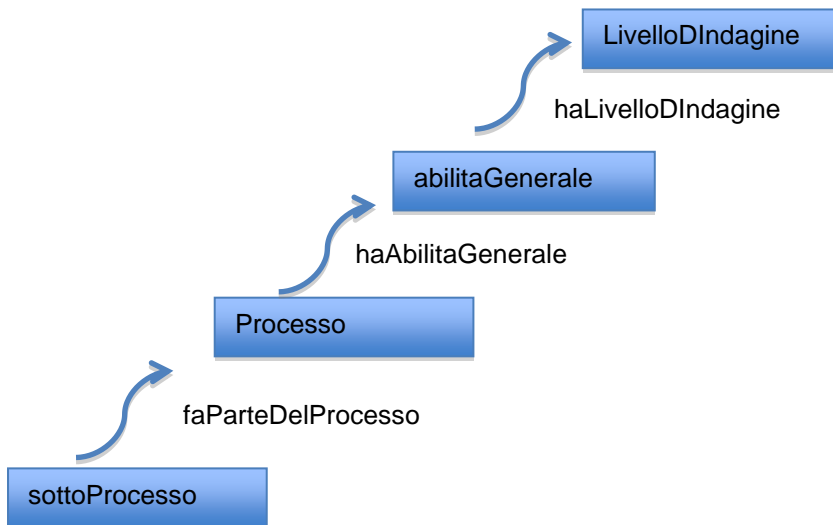
haFormato, **haLingua** e **haTipo** sono tre relazioni che associano ad ogni learning object il suo formato, la lingua in cui è scritto ed il suo tipo.

haProcessoOSottoprocesso è la relazione che associa ad ogni learning object i processi e/o i sotto-processi formativi sui quali si focalizza tale learning object. È da notare che un learning object potrebbe essere associato sia con uno o più processi che con uno o più sottoprocessi.



Il secondo gruppo di relazioni viene utilizzato per rappresentare la struttura gerarchica dei livelli di indagine. Ogni elemento più basso della gerarchia è

collegato al proprio padre con una relazione specifica secondo lo schema seguente



faParteDelProcesso è la relazione che associa ad ogni sottoprocesso il processo a cui appartiene. Ad esempio il sottoprocesso *Letto-scrittura-Decodifica-Lettere-Riconoscimento*, è associato al processo *Letto-scrittura-Decodifica-lettere*.

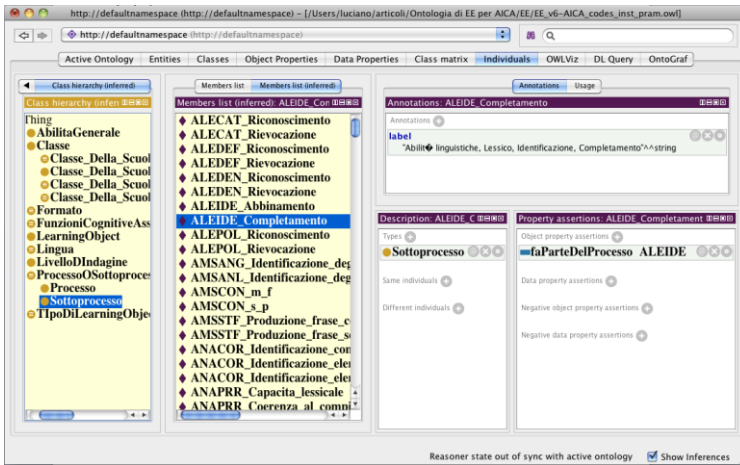
haAbilitaGenerale è la relazione che collega un processo all'abilità generale a cui si riferisce. Per esempio il processo *metafonologia-globale-rime*, è collegato tramite la relazione *haAbilitaGenerale* all'abilità generale *metafonologia-generale*; e così sarà anche per il processo *Metafonologia-globale-sillabe*; mentre il processo *metafonologia-profonda-fonema* sarà collegato da questa relazione all'abilità generale *metafonologia profonda*.

haLivelloDIndagine è la relazione che associa ad ogni abilità generale il livello di indagine a cui fa riferimento. Per esempio l'abilità generale *Letto-scrittura-codifica*, è associata tramite la relazione *haLivelloDIndagine* al livello di indagine *Letto-scrittura*.

Un'ultima relazione che non è riconducibile a nessuna dei due gruppi precedenti è la relazione **haClassePrecedente** che rappresenta l'ordine temporale fra le diverse classi. In particolare questa relazione collega ogni classe, ad esclusione della prima classe della scuola dell'infanzia, con la classe precedente.

La A-box (*assertional Box*) dell'ontologia contiene tutti gli oggetti con la loro caratterizzazione di tipo (cioè la loro appartenenza a una o più classi) e le specifiche relazioni tra gli oggetti. Nella figura seguente viene mostrato uno screenshot di *Protegé* che mostra nel riquadro centrale tutti gli oggetti della classe *sottoprocesso* (evidenziato in blu nella gerarchia delle classi, prima finestra a sinistra) e si focalizza sul processo *Abilità linguistiche, Lessico, Identificazione, Completamento* (evidenziato in blu nella seconda finestra),

mostrando come questo sotto processo sia collegato con il processo relazione *faParteDelProcesso ALEIDE (Abilità linguistiche, Lessico, Identificazione)* tramite la relazione *faParteDelProcesso (ultima finestra a destra)*



3.3 Esempi di *query* all'ontologia di E-VALUE

Seguono due esempi di *query* in linguaggio SPARQL (<http://www.w3.org/TR/rdf-sparql-query/>) utilizzato nella prima fase della progettazione del DSS (descritto nella sezione 3.2 come "selezionare tra tutti i *learning object* disponibili nel catalogo Erickson quelli rilevanti per i processi/sottoprocessi specificati nell'input").

Esempio 1. Seleziona tutte le coppie di *processi amici* in cui il primo processo sia classificato come *Matematica-Geometria*

```
SELECT ?sp ?spa
WHERE {
  ?sp :haProcessoAmico ?spa.
  ?sp :faParteDelProcesso ?p.
  ?p :haAbilitaGenerale ?ag.
  ?ag :haLivelloDIndagine :Matematica-Geometria
}
```

Esempio 2. Seleziona tutti i *learning object* del tipo *Libro*, scritti in italiano e classificati *MMASNL (Metafonologia, Manipolazione di suoni non linguistici)*

```
SELECT ?lo
WHERE {
  ?lo :haProcessoOSottoprocesso ?sp.
  ?sp :faParteDelProcesso ?p.
  ?p :haAbilitaGenerale :MMASNL.
  ?lo :haLingua :ITA.
```

```
    ?lo :haTipo :Libro  
}
```

5. Studi correlati

È stata svolta un'analisi sulle forme di catalogazione di materiale digitale utilizzate con maggiore frequenza (ONIX, Dublin Core, gli standard catalografici di SCORM-LOM) dalla quale è emersa la necessità di non limitarsi a scegliere uno di questi sistemi ma di crearne uno che fosse sufficientemente elastico da consentire una catalogazione dettagliata secondo i punti di maggiore interesse per il progetto E-VALUE e con una struttura parzialmente a faccette.

6. Conclusioni

L'ontologia qui presentata intende descrivere e catalogare i materiali che le Edizioni Centro Studi Erickson utilizzano per il progetto E-VALUE, e può essere considerata come un punto di partenza per la creazione di un DSS a scopo formativo.

Un'osservazione importante è che il processo di creazione dell'ontologia non si limita alla traduzione automatica o semiautomatica della classificazione in linguaggio OWL. Si è verificato invece un complesso processo interattivo in cui il lavoro di classificazione è stato verificato e parzialmente rivisto per permettere la rappresentazione dello schema catalografico in linguaggio logico/formale.

Inoltre la fase di creazione dell'ontologia per il progetto E-VALUE si è rivelata anche una "fase di studio" che ha permesso alla parte tecnica del progetto (Fondazione Bruno Kessler) di approfondire la conoscenza in questo settore, e agli esperti del settore (Edizioni Centro Studi Erickson) di avvicinarsi alle logiche della modellazione formale.

7. Bibliografia

AID – Associazione Italiana Dislessia, Comitato Promotore Consensus Conference (a cura di) *Disturbi Evolutivi Specifici di Apprendimento. Raccomandazioni per la pratica clinica di dislessia, disortografia, disgrafia e discalculia*, Trento, Erickson, 2009.

Baader F. [et al.] (eds.), *The Description Logic Handbook: Theory, Implementation, and Applications*, Cambridge University Press, Cambridge MT, 2003

Gnoli, C., *Classificazione a faccette*, AIB, Roma, 2004

Gnoli C., Marino V., Rosati L., *Organizzare la conoscenza. Dalle biblioteche all'architettura dell'informazione per il Web*, Tecniche Nuove, Milano, 2006

Gruber, T. R., *A translation approach to portable ontologies*. *Knowledge Acquisition*, 5(2):199-220, 1993

"OWL 2 Web Ontology Language Document Overview". W3C. 2009-10-27, 2009.